

DARKTRACE

CISO向け AI導入ガイド

投資の前に、まず見極めを：
AI導入への第一歩を踏み出す

Secure
What AI
Can Do

コンテンツ

02	はじめに
03	AIへの投資の際に検討すべきこと
03	AISPを評価する前に知っておくべきこと
04	サイバーセキュリティにおけるAIの限界の理解
05	AIベンダーの評価方法
06	ガバナンス、安全性、データ管理
07	データの収集とトレーニング
07	モデルと手法の選定
08	パフォーマンスと精度の検証
08	解釈可能性、調整可能性、透明性
09	AIサイバーセキュリティベンダーとしてのDarktrace
10	結論

はじめに

AIは急速にセキュリティ製品に組み込まれており、今後も、これらの製品による脅威の検知、分析、対応において大きな役割を果たし続けます。時間とリソースを節約し、価値あるデータからインテリジェンスとインサイトを抽出し、従来のセキュリティソリューションよりも詳細かつ複雑な分析を実行することで、AIは組織の能力とセキュリティ体制を強化します。

AI導入の加速に加え、近年のエージェント型AIと自律型エージェントの急速な普及に伴い、CISOはこうした機能の表面だけを見るのではなく、異なるモデルがどのように構築、トレーニング、検証されているか、そして新しいデータや未知のデータに直面した際にどのように動作するかを理解する必要があります。

ベンダーは幅広いアプローチを採用しており、それぞれの設計上の選択が、精度、信頼性、そしてブラインドスポットを生む可能性に影響します。AI機能がどのように動作するか、また、解決するとされる課題に本当に適しているかを明確に把握することで、セキュリティリーダーはより的確な質問を行い、不必要なリスクを加えることなく、そのテクノロジーが適切に活用されていることを確認できます。

同様に重要なのは、AIが安全に導入され、責任を持って管理されているかを把握することです。可視性の低さ、不十分な制御、一貫性のないメンテナンスは、パフォーマンスの低下、モデルドリフト、バイアス、悪用可能な脆弱性につながる可能性があります。

適切に設計されたソリューションには、強固なセーフガード、透明性の高い推論プロセス、データの取り扱いと保護に関する明確な制御機能が備わっています。CISOは、ベンダーのガバナンスフレームワーク、テストと検証のワークフロー、悪意のある入力に対するレジリエンス、不正使用やデータ漏洩を防ぐセーフガードについて、自信を持ってベンダーに質問できる必要があります。近年のAI投資、開発、ベンダーによる主張の急増を考慮すると、このようなレベルのデューデリジェンスは特に重要です。

AIは生産性、洞察力、セキュリティ効果において大きな改善をもたらす可能性を秘めている一方で、その市場は誇大広告、誤解、透明性の欠如によっても形成されています。

本ガイドは、CISOがそうした雑音を排し、実証済みの能力と流行語を見分けられるよう設計されています。これにより、AIを活用したセキュリティツールを自信を持って評価し、導入が信頼でき、責任あるものであり、



AIへの投資の際に検討すべきこと

AIサービスプロバイダー（AISP）とは、AI-as-a-Serviceを構築・提供するベンダーです。モデルの選定、設計、学習から、改良、検証、継続的なガバナンスまで、すべてを担います。AISPのソリューションは、単一のモデルで構成される場合もあれば、複数のモデルを組み合わせて配置し、より複雑で包括的な出力を生成する場合もあります。

ほぼすべてのセキュリティベンダーが、自社ソリューションの一部としてAIの活用を始めています。サイバーセキュリティにおいて、AIは非常に価値の高いものになり得ます。従来のセキュリティシステムはシグネチャベースの検知に依存していますが、AIシステムはパターンや傾向を特定し、汎用化によって新たな攻撃を検知することができます。

AISPを評価する前に知っておくべきこと



AISPは単一のソリューション内で複数の異なるモデルを使用できる



AIは幅広い技術を網羅しています（生成AIを含みますが、これに限定されません）



AIは常に正しいとは限らず、最適化が必要



すべてのAIシステムには固有のリスクが伴うため、ガバナンスが必要

AIは常に進化しており、「AIシステムとは何か」の明確な定義を維持することが難しくなっています。経済協力開発機構（OECD）は、AIを、明示的なプログラミングを必要とし、入力から出力を推測できる機械ベースのシステムと定義しています。

AIの分野は広範であり、クラスタリング、分類、回帰、コンピュータビジョン、大規模言語モデル（LLM）、ドメイン固有言語モデル（DSLM）、エージェンティックAIなど、幅広い技術がAIの傘下に含まれます。

LLMとDSLM

近年急速に普及しているAI分野の1つが、大規模言語モデル（LLM）です。LLMは自由形式のテキスト入力を処理するように設計された汎用モデルであり、自然言語タスクに適しています。

同様に、ドメイン固有言語モデル（DSLM）は、特定のタスク向け、または特定のドメインで機能するように設計されたLLMであり、セキュリティ関連のタスク向けに特化してトレーニングされたDarktraceのDEMIST-2言語モデルなどがその一例です。

生成AI

生成AIとは、自然言語テキスト、画像、音声などの独自の出力を生成できるAIシステムを指します。こうしたシステムは、LLMを基盤とするクラウドベースのAIチャットボットシステムの台頭に伴い、急速に普及しています。生成AIはLLMのみを指すわけではありません。LLMと同じ技術やその他の生成型統計手法に基づいて構築された画像や動画の生成モデル（OpenAIのDalle-2やGoogleのVeoなど）のように、他にも多くの生成AIシステムが存在します。

また、LLMは必ずしも生成用途で使用する必要はないという点にも留意することが重要です。一部のタスクにおいては、生成AIよりも適している他の幅広いLLMベースのアプローチが存在します。

エージェント型AI

一部のAISPは最近、エージェント型AIソリューションの提供に大きく注力しています。「エージェント型」は現在AI業界で頻繁に使われている流行語であり、さまざまな定義が飛び交う中、多くのAISPが自社システムをエージェント型であると主張しています。

エージェント型AIの定義はベンダーによって大きく異なり、あらゆる自律型システムから、ツールを利用可能な生成AIシステム、さらにはあらゆるAIシステムに適用できる単なるマーケティング用語に至るまで多岐にわたります。

このようなばらつきがあるため、導入を検討しているAISP（AIサービスプロバイダー）に対して、エージェント型AIをどのように定義しているか、自社のソリューションにエージェント型AIがどのように実装されているかを確認し、その実装によってどのようなリスクが生じる可能性があるかを検討する必要があります。

複合AI

AISPが活用できるもう1つの主要な技術として、複合AIが挙げられます。これは、複数の異なるAI技術を単一のソリューションに組み合わせたものです。これにより、複数のシステムが相互に補完し合うことで死角を減らし、ソリューションの適用範囲を拡張できます。

サイバーセキュリティにおけるAIの限界の理解

ソリューションにAIシステムを導入することで、独自のインサイトが得られ、自律的なアクションによって効率化が図られるというメリットがある一方で、すべてのAIシステムには固有のデメリットや潜在的なリスクが存在することを忘れてはなりません。

例えば、生成AIシステムはその柔軟性や、非構造化自然言語データを処理・出力する能力によってますます普及しているものの、タスクによって得意不得意があります。

AIソリューションへの投資を検討する際、購入担当者は、AIが他のアプローチと比べてどのような利点をもたらすかを評価し、利用可能な多様なAI手法と、それぞれが提供する独自のメリットを検討する必要があります。本ペーパーでは、AIの開発、実装、ガバナンスに対するAISP（AIサービスプロバイダー）のアプローチを理解するために、尋ねるべき主な質問をいくつか取り上げます。

最後に、AIシステムの導入が急速に加速する一方で、それらを取り巻くガバナンスやリスク管理の多くは依然として初期段階にあります。さらに、システムが複雑化するにつれて、ガバナンスやリスク管理に関する要件は増大する一方です。ベンダーがどのような制御機能を備えているかを理解することは、印象的な技術デモと、信頼性が高く堅牢な企業利用が可能なシステムを区別する上で重要な要素となり得ます。

特に生成AIに関して、Gartnerは次のように指摘しています。



コンテンツ生成、対話型ユーザーインターフェース、ナレッジディスカバリーに極めて有用



セグメンテーション/分類、レコメンデーションシステム、知覚、インテリジェントオートメーション、異常検知/監視、自律システムなどに多少有用



予測/フォーカシング、計画、意思決定インテリジェンスにはほとんど役に立たない

また、Gartnerの一部のリサーチでは、次のような結論に至っています。

より高精度で堅牢なAIソリューションを実現するためには、AIリーダーは単一のモデルや手法の利用にとどまらず、コンポジットAIの実践を取り入れ、AIシステム全体を包括的に捉える視点を持つ必要があります。

AIベンダーの評価方法

AISPの初期調査において、購入検討者が候補を絞り込むために確認できるポイントがいくつかあります。これは、証拠となる事例や顧客の声を探すことで確認できます。

これらは、そのAIツールが有効に機能し、貴社でも同様に効果を発揮することを示す証拠となります。特に、業界、組織規模、地域、その他の属性など、貴社と同様の状況にあるお客様からのインサイトに着目してください。

また、AISP（AIサービスプロバイダー）を評価する際には、そのプロバイダーに優れたガバナンス文化が備わっているかを評価することも重要です。これは、組織がAIを構築、テスト、展開する方法に、ガバナンスがどのように組み込まれているかを理解することを意味します。

弊社のDarktrace「AIサイバーセキュリティの現状」レポートの調査によると、AIガバナンスポリシーを策定している組織はわずか37%にとどまることが明らかになりました。



企業がAIガバナンスポリシーを策定していたり、既知のAIガバナンス標準に準拠していると主張していたりしても、その企業がそれらのポリシーやプロセスを実際にどのように運用しているかを問うことが重要であることを認識する必要があります。企業におけるAIガバナンスの文化と定着度が高ければ高いほど、それがその企業の製品に反映される可能性も高まります。

導入検討者は、選定候補となるAISP（AIサービスプロバイダー）に対し、実施されているガバナンスやデータセキュリティのプロセス、使用しているテクノロジー、テスト・評価・妥当性確認・検証（TEVV）のために従っているワークフローについて質問する必要があります。これらの質問を行うことで、導入検討者は、ソリューションが自社のビジネス要件やポリシーを満たしているか確認するための詳細な情報を収集できます。

以下のセクションでは、購入担当者が5つの主要カテゴリーについて尋ねることができる質問を概説します。

- 01 ガバナンス、安全性、データ管理
- 02 データの収集とトレーニング
- 03 モデルと手法の選定
- 04 パフォーマンスと精度の検証
- 05 解釈可能性、調整可能性、透明性



ガバナンス、安全性、データ管理

AIガバナンスは、ソリューション内のすべてのAIシステムとモデルが安全、高精度、かつ適切に維持されていることを保証するだけでなく、システムに渡されるすべてのデータが責任ある倫理的かつ法に準拠した方法で使用されていることを検証するためにも不可欠です。強固なガバナンスシステムは、内部のすべてのシステムに対して十分な可視性を提供し、組織がデータの安全性、セキュリティ、完全性を確保できるようにします。

多数のモデルを備える、より複雑な複合AIシステムでは、ガバナンスが不可欠です。複数のAIシステムを備えるソリューションでは、すべてのシステムを追跡・文書化する十分に堅牢なガバナンスソリューションがなければ、すべてのモデルを最新かつ安全に保ち、精度向上のために継続的に改善することは困難です。

社内システムに対する一貫性のないメンテナンスや可視性の欠如は、脆弱性や弱点、バイアスの混入につながり、ソリューションのパフォーマンスや精度に悪影響を及ぼす可能性があります。

以下は、AIサービスプロバイダー候補のガバナンスポリシーやプロセスに関する情報を収集する際に、バイヤーが尋ねるべき質問の例です。

ベンダーへの質問事項

どのようなAIガバナンスポリシーやフレームワークに従っていますか？ また、現在どのような認証を取得していますか？

AIソリューションにおいて、クライアント情報（IP、PIIなど）のデータセキュリティ、プライバシー、リングフェンシングをどのように確保していますか？

悪意のある入力、データポイズニング、その他の敵対的攻撃からAIシステムをどのように保護しますか？

顧客または自社のデータを処理または取り扱うために、外部でホストされているサードパーティのモデルやプラットフォームを利用していますか？ その場合、これらのシステムの安全性、セキュリティ、機密性をどのように確保していますか？

ガバナンスやデータセキュリティに関する質問への回答を確認する際、購入者は、利用するAISP（AIサービスプロバイダー）が強固なガバナンスを確保するために採用している具体的な認証、ワークフロー、またはフレームワークへの言及があるかを確認する必要があります。



DarktraceはISO/IEC 42001規格の認証を取得しています。ISO/IEC 42001は、AIがもたらす特有の課題に関する貴重なガイダンスを提供する、世界初のAIマネジメントシステム（AIMS）規格です。組織にとって、この規格はAIに伴うリスクと機会を体系的に管理する方法を提示し、あらゆるAIシステムの強固で持続的なガバナンスを支援します。

利用可能なAIガバナンスフレームワークはISO/IEC 42001だけでなく、NIST AIリスク管理フレームワークのような他のフレームワークも広く利用されている点に留意することが重要です。しかし、ISO/IEC 42001は自己申告による準拠に依存するのではなく、外部監査を通じてさらなるレベルの保証を提供します。

AIガバナンスポリシーに加え、データセキュリティとプライバシーはあらゆるシステムにおいて極めて重要です。そのため、データセキュリティフレームワークや認証を確認することは、システムが十分に安全であり、データを安全に取り扱うことができるという安心感をバイヤーに与えることにもつながります。

データの収集とトレーニング

AIシステムを構築する上で最も重要なステップの1つは、学習データの初期収集です。低品質、一貫性の欠如、不正確、または偏りのあるデータでAIシステムを学習させた場合、構築されたシステムが最適なパフォーマンスを発揮することは決してありません。

ISO/IEC 42001などのガバナンスフレームワークは、AIシステム向けのデータ収集時に準拠すべき管理策を定めています。これには、学習データの追跡可能性と改ざんされていないことを担保するためのデータ元の特定や監査ログの記録、リスクの状況に応じた収集ユーザーデータの十分な非特定化や抽象化の実施、必要に応じたバイアスの徹底的な検証などが含まれます。

以下は、データ収集に関する質問を検討する際の出発点として、バイヤーが参照できる質問のリストです。

ベンダーへの質問事項

AIモデルや学習データのバイアスを防ぐために、どのような対策を講じていますか？

学習データに関連するリスクをどのように最小化しますか（例：反芻、データの出所）？

回答を精査する際、バイヤーは、プロバイダーが学習データの調達や取り扱いを管理するAIマネジメントシステムを整備しているかどうか、そしてバイアスや体系的な問題を特定・軽減するための明確なプロセスが存在するかどうかを評価する必要があります。

AIの学習データに個人識別情報（PII）が含まれる場合（組織固有の自己学習システムを除く）、AIサービスプロバイダー（AISP）は、必要に応じてデータの非識別化（PIIの削除）に使用される具体的な管理策やプロセス、またはその他のあらゆる匿名化プロセスを提示する必要があります。

モデルと手法の選定

タスクによって、適したAIモデルや手法は異なります。特定のタスクに適したモデルを選択することは、ソリューションの正確性と効率性を確保するための重要なステップです。

モデルや技術にも強力なガバナンスが必要であり、それは単に規格や認証を遵守することだけではなく、意図した目的を達成するために最も効果的かつ効率的なモデルやプロセスを採用することでもあります。これには、使用されているモデルや手法の詳細なログを記録し、サードパーティ製モデルや社内モデルのサプライチェーンリスクを含め、リスクや脆弱性がないか継続的に分析することが含まれます。

AIサービスプロバイダーが内部ホスト型および外部ホスト型のモデルやサービスを選定する際に適切なガバナンス手順に従っていることを確認するために、購入者が確認すべき質問事項を以下に示します。

ベンダーへの質問事項

ソリューションではどのような種類のAIモデルを利用していますか？

生成AI（社内または社外でホストされているもの）を利用していますか？また、利用している場合、それらに対してどのような保護策やガードレールを講じていますか？

それぞれのタスクに最適なモデルを選択していることをどのように保証しますか？

エージェンティックAIを利用していますか？利用している場合、どのように、どこで利用されていますか？

基礎モデルと汎用モデル、どちらを使用していますか？使用している場合、これらはどのようにホスティングされていますか？

AISPが使用するカスタムシステムに関する具体的な詳細情報は独自のものである可能性が高いものの、バイヤーはベンダーが採用している広範な技術の概要を提供できることを期待すべきです。サードパーティ製またはクラウドホスト型の基盤モデル、関連するAPI、該当するデータレジデンシーの利用に関する透明性は、データの取り扱いと完全性に関する確証を求める購入者にとって、依然として重要な検討事項です。

安全性とセキュリティを確保するため、AISPは外部システムを導入する前に、サードパーティに対するデューデリジェンスや徹底した調査・分析を参照する必要があります。

パフォーマンス と精度の検証

AIシステムは、一般的な利用を開始する前に、詳細なTEVV（テスト、評価、検証、妥当性確認）チェックを受ける必要があります。システムの評価と、それに続く最適化、調整、改善は、構築されるシステムが適切かつ安全に調整され、意図したタスクに対して有意義なインサイトを提供できるようにするために不可欠です

TEVVには、単なる精度の検証にとどまらない多くの要素が含まれます。システムが偏りがなく、データに過剰適合したり過小適合したりせず、効率的かつ正確に動作することを確認するためには、多くの手法を適用する必要があります。また、望ましくないデータドリフトやモデルドリフトを防ぐためには、リリース後の継続的な評価もきわめて重要です。テストではモデルの物理的なパフォーマンスも対象にし、適切な時間枠で出力を生成できること、そして妥当なリソース消費で動作することを確認する必要があります。

ISO/IEC 42001では、すべてのAIシステムを本番環境で使用する前に、システムの関連性と正確性を評価するワークフローを作成し、完了することが義務付けられています。

以下は、AIサービスプロバイダーが適切なTEVVを実施しているか、またパフォーマンスと精度の検証において十分なガバナンスが確保されているかを確認するために、購入者が尋ねることができる質問のリストです。

ベンダーへの質問事項

AIモデルの出力をどのように監査、テスト、評価、確認、検証していますか？

モデルは絶えず変化する外部の脅威ランドスケープにどのように適応し、データドリフトのリスクをどのように低減しますか？

TEVVワークフローには、レイテンシ、メモリ消費量、エネルギーおよびリソース要件など、物理的パフォーマンスのテストが含まれていますか？

これらの質問に対する回答には、既存のAIMS（またはその他のガバナンスシステム）の一部であるか個別のフレームワークであるかを問わず、TEVVワークフローへの直接的な言及が含まれる必要があります。

解釈可能性、 調整可能性、透明性

AIシステムにおいて重要なのは、データ出力が解釈可能であること、その出力に至るまでの手順が理解可能であること、そしてエンドユーザーがAIの出力が正しいと信頼できるようにすることです。AIシステムはしばしばブラックボックスですが、正確な手順が隠されている場合でも、ユーザーはシステムが特定の結論に至った理由を理解する必要があります。

サイバーセキュリティの世界では、AIシステムへの信頼が不可欠です。これらのシステムはリアルタイムでアクションを実行し、接続の遮断、メッセージの拒否、ネットワークアクセスの制限などを自律的に行う必要があります。システムがこれらのアクションを正確かつ適切に実行できることが信頼できなければ、エンドユーザーが機能を無効化したり、アクションを取り消したりする傾向が生じ、システムの防御機能が大幅に低下するおそれがあります。サイバーセキュリティにおいて極めて高い信頼性が求められることから、そのユースケースの多くは生成AIによる自動化にはあまり適していません。AIシステムの調整機能は重要な要素です。

組織ごとに、行動、構造、文化は異なります。AIモデルをファインチューニングする能力は、全体的な目標の達成に役立ち、より幅広いユースケースに適したソリューションを実現します。

以下の質問は、システムが十分に解釈可能かつ調整可能であり、人間のアナリストとAIツールの間に強固な信頼関係を促進できることを確認するために、バイヤーがAISPに尋ねる質問の出発点となります。

ベンダーへの質問事項

AIシステムの解釈可能性をどのように確保していますか？

人間のアナリストとAIの出力結果との間に信頼関係を築くにはどうすればよいでしょうか？

SOCチームは、独自のポリシーやビジネス要件に合わせてモデルをどのようにファインチューニングできますか？

AISPは、自己学習と自律的な適応、またはエンドユーザーが利用できる手動調整を通じて、組織固有の構造、ポリシー、ニーズに合わせてソリューションをどのように適応できるかを説明する必要があります。

透明性と解釈可能性が考慮されていることの確認も重要です。AISPは、完全な透明性を確保し、自社のソリューションに対する信頼を促進するためのポリシー、ワークフロー、手法に関する情報を提供する必要があります。

AIサイバーセキュリティベンダーとしてのDarktrace

Darktraceは10年以上にわたりサイバーセキュリティにおけるAIの構築と適用に取り組み、ますます複雑化し急速に変化する脅威環境とともに、その機能を発展させてきました。

この経験により、AIに対する成熟した多層的なアプローチが実現しました。DarktraceのSelf-Learning AIは、事前定義されたルールや既知の攻撃シグネチャに依存することなく、各組織の正常なパターンを継続的に学習して行動を理解し、コンテキストを解釈し、重大な逸脱を特定します。これにより、時間をかけて培われた実証済みの行動理解が可能になり、他では見逃されがちなリスクのわずかな兆候を明らかにするのに役立ちます。

このアプローチは、ネットワーク、ID、クラウド、メール、アプリケーションにまたがるアクティビティが発生する、動的なクロスドメイン環境で開発・検証されてきました。システム間を移動する挙動を理解することで、Darktraceは組織全体のアクティビティを継続的に可視化し、環境や脅威の変化に応じて適応し続ける検知、調査、対応を可能にします。これは、実環境での適用と改良に基づく長年にわたるイノベーションへの取り組みと、責任ある透明性の高いAI開発への注力を反映しています。

この分野のパイオニアとして、当社は責任あるアプローチをプロセスの基本原則の中核に据えることにも注力してきました。この取り組みはISO/IEC 42001の認証取得にも反映されており、サイバーセキュリティのレジリエンス強化に向けて、業界をリードするSelf-Learning AIを提供するDarktraceの実力を実証しています。

当社のステークホルダー、顧客、パートナーの皆様は、Darktraceが責任ある倫理的かつ安全な方法でAIシステムを開発しており、日々の業務におけるAIの利用を法令遵守、安全かつ倫理的な方法で管理していることを確信していただけます。

DarktraceがAIシステムを保護する仕組み

Darktraceは、Darktrace / SECURE AIによって、組織全体のAIシステムから発生するリスクの監視・対応機能を新たに提供します。このソリューションは、各組織のコンテキストにおいてプロンプト、エージェント、システムがどのように使用されているかを分析し、AIとのあらゆる相互動作を単一のビューに集約します。この独自のアプローチにより、チームは人間とAIエージェントの両方のアクティビティに対して意図を理解し、リスクを評価し、機密データを保護し、ポリシーを適用することができます。生成AI、AIエージェント、開発環境、シャドーAI全体にわたるリアルタイムの可視性と制御を提供し、ルールベースのアプローチでは見逃されがちな不正利用、設定ミス、ドリフトをチームが特定できるようにします。

ユーザーやシステム全体の文脈と意図を理解することで、企業全体でAIの導入が拡大する中、組織が監視体制を維持できるよう支援します。

[Secure AI Readiness Programへの参加はこちらから](#)：AI脅威の最新ニュース、AIセキュリティを取り巻く新たなアプローチ、この分野におけるDarktraceの開発情報を含めた最新イノベーションに関する考察をお届けします。

AIの保護についてDarktraceのエキスパートにご相談をご希望ですか？

お客様のビジネスにとって最も問題となるAIリスクについて、ガバナンス、可視性、リスク削減、長期的な準備など、どの部分に集中すべきかを含め、実務的な情報をご提供します。



サイバーセキュリティにおける責任あるAIに対するDarktraceのアプローチ

学界や政府機関のグローバルな専門家との連携に基づき、説明責任、説明可能性、そして継続的な検証がサイバーセキュリティ技術にどのように組み込まれているかを詳述した、Darktraceの責任あるAIアプローチの原則をご覧ください。

詳細は[こちら](#)

AIの武器庫

ホワイトペーパー『AIの武器庫』で、教師あり機械学習、教師なし機械学習、そしてLLMをサイバーセキュリティにどのように適用できるか、またDarktraceが特定の技術を多層的なアーキテクチャに統合して人間のセキュリティチームをどのように強化しているかについて、理解を深めることができます。

詳細は[こちら](#)



結論

AIがサイバーセキュリティにますます深く組み込まれるにつれ、セキュリティリーダーにとって最も重要な要素は、製品がAIを使用しているかどうかではなく、そのAIがどのように設計され、管理され、検証され、制御されているかとなります。購入者は、データの調達元と保護、モデル選定、テストと検証の実践、敵対的入力に対するレジリエンス、AI主導の意思決定の透明性について、明確な回答を期待すべきです。

強力なAIソリューションは、慎重に選定された技術、継続的な評価、そして時間の経過とともにリスクを低減するガバナンスフレームワークを組み合わせて構築されます。これにより、脅威が進化しても、システムの正確性、適応性、組織のポリシーとの整合性が維持されます。

責任を持って導入すれば、AIは大きなメリットをもたらします。例えば、新たな脅威の検出精度の向上、より迅速かつ正確な調査、セキュリティチームの運用負担の軽減、そして自律的な対応に対する信頼性の向上などが挙げられます。Darktraceの多層的な自己学習アプローチは、多様なAI技術、強固なガバナンス、そして説明可能な成果が連携することで、信頼性の高いサイバーセキュリティを大規模に実現できることを実証しています。

個別デモについては Darktraceまでお問い合わせください

詳細を見る ➔

■ Darktraceについて

Darktraceは、組織が日々変化する脅威環境に先手を打てるよう支援する、AIサイバーセキュリティのグローバルリーダーです。2013年に英国ケンブリッジで設立されたDarktraceは、各企業からリアルタイムで学習するAIを活用し、未知の脅威から組織を守るための不可欠なサイバーセキュリティプラットフォームを提供しています。Darktraceのプラットフォームとサービスは、世界中の約10,000社のお客様を保護する2,700人以上の従業員によって支えられています。詳細については、www.darktrace.comをご覧ください。